EXHIBIT 2

Diacetyl as a Flavor Component in Full Fat Cottage Cheese

M.J. ANTINONE, H.T. LAWLESS, R.A. LEDFORD and M. JOHNSTON

ABSTRACT

Diacetyl was detected at about 0.2 ppm using a two alternative forced choice procedure. Consumer acceptance for cottage cheese showed a small effect of added diacetyl, peaking at 1.0 ppm for aroma acceptance and flavor acceptance. Patterns of acceptance across diacetyl levels were similar for both frequent and infrequent consumers. Consumer acceptance tests conducted in a campus-based sensory testing facility gave similar results to tests conducted in a shopping mall, although the mall data were more variable.

Key Words: diacetyl, cottage cheese, flavor, thresholds, sensory evaluation

INTRODUCTION

DIACETYL is a four carbon dicarbonyl molecule that imparts a characteristic buttery aroma and flavor to cultured dairy products (Antinone, 1990; Collins, 1972; Parker and Elliker, 1953; Veringa et al., 1984). Most of the diacetyl present in dairy products is formed from fermentation of citric acid by the starter culture organisms used to manufacture the product (Collins, 1972). Diacetyl in dairy products can also result from contamination by diacetyl-producing yeasts (Collins, 1972) and bacteria such as certain *Pseudomonas* species (Antinone, 1990; Sebek, 1952), or through direct addition of diacetyl to the product. Diacetyl in the form of starter culture distillate may be added to cottage cheese dressing to produce a more uniform flavor profile across batches (Ernstrom, 1965).

The amount of diacetyl in cultured dairy foods varies from product to product. Veringa et al. (1984) reported that the diacetyl content of many cheeses, including cottage cheese was 1 to 2 ppm. However six of nine different brands of New York State cottage cheese contained < 0.62 ppm diacetyl (Antinone, 1990). Parker and Elliker (1953) proposed that cottage cheeses that contain less than 0.62 ppm of diacetyl have a flat aroma. Furthermore, they reported a relationship between decreasing diacetyl content and decreasing flavor scores. When cottage cheese samples with diacetyl concentrations between 1.2 and 3.2 ppm were rated for flavor, the samples which contained 1.2 ppm diacetyl were given the lowest flavor scores. Six of the nine New York State samples indicated above would have been expected to have a flat aroma and lower flavor scores under these criteria. Parker and Elliker's study used four experienced dairy product judges so their data may not be representative of the consuming public. The number of judges was too small to make statistical inferences. Furthermore, trained judges may be better than consumers at detecting and discriminating diacetyl levels in cottage cheese and may be more critical. McBride and Hall (1979) and Schroeder et al. (1988) showed that trained dairy product judges' scores of cheddar cheese did not always correlate well with consumer acceptance.

Currently, the consumer appeal of diacetyl flavor in cottage cheese is poorly understood. In a progress report, Geilman (1992) found that female consumers preferred cottage cheese that contained 2 ppm of diacetyl over samples containing 0, 1, or 6 ppm of diacetyl. Results from our laboratory suggest that there is a high variability in sensitivity to diacetyl among consumers (Lawless et al., in press). The objectives of our current experiment were to determine if consumers can detect levels of diacetyl that are commonly found in cottage cheese and to determine an optimal consumer acceptance level for diacetyl.

The authors are with the Dept. of Food Science, Cornell Univ., Ithaca, NY 14853. Address inquiries to Dr. H.T. Lawless.

MATERIALS & METHODS

Sensory evaluation procedures

Diacetyl detection. Cottage cheese samples had diacetyl added 24 hr before each testing session in ≈ 5L plastic containers with lids. Samples were stored (4°C) in the dark until the day of testing. Samples (about 25g) were transferred to ≈120mL plastic cups with lids numbered with 3-digit random codes on the day of testing. Testing was carried out from 11:30 am to 4:30 pm. Concentrations of diacetyl increased in binary steps as follows: 0.125 to 32.0, 0.0625 to 32.0, and 0.0625 to 8.0 ppm for the triangle, three alternative forced choice (3-AFC) and two alternative forced choice (2-AFC) tests, respectively. All samples were presented in ascending order with randomization within each pair or triad.

For the triangle test, panelists were instructed to select the sample which was most different from the other two. No training was given and panelists were not instructed as to the nature of the difference under study. Half the triads contained two samples with added diacetyl at each diacetyl concentration level and one sample with no added diacetyl, while the other half of the triads contained one with added and two nonadded samples. Panelists were told that the differences between samples were due to addition of a "flavor additive" and that they were to select the sample that possessed the most noticeable flavor. Panelists were not told what the flavor was.

Due to unexpectedly low levels of discrimination on the triangle and 3-AFC tests, the simpler and possibly more sensitive 2-AFC test was performed. The 2-AFC test had one sample that had added diacetyl and one sample that did not in each pair. Panelists were told that the samples differed in diacetyl level and were instructed to select the sample that contained the most diacetyl. A 30 min training session was held before each week's test to familiarize panelists with diacetyl aroma and flavor and the nature of the 2-AFC test.

Both the triangle and 3-AFC tests were performed once over four testing days. Each panelist received an average of two triads per day. The 2-AFC test was performed in duplicate over five days and then repeated the following week with a new lot of cottage cheese for a total of four replications. The panels consisted of 50 individuals for the triangle and 45 for 3-AFC tests. The 2-AFC test consisted of 20 panelists, 10 who performed above average and 10 who performed below average on the 3-AFC tests. Above average panelists identified the diacetyl sample correctly at least 80% of the time, whereas below average panelists were correct less than 51% of the time. One panelist who was included in the above average group did not participate in the 3-AFC test but performed 67% of the triads correctly on the triangle test. These groups were tested to see if individual differences in ability to detect diacetyl in cottage cheese were stable characteristics. Panelists were healthy adults selected from the Cornell University community ranging in age from 18 to 59 and had no self-reported problems in their sense of smell or taste. Panelists were given crackers and spring water and instructed to chew a piece of cracker and rinse after each sample to minimize carry-over and adaptation effects.

Consumer acceptance tests. A total of 200 consumers were recruited at two different locations on two consecutive days. Consumers ranged in age from 12 to 66 and had no self-reported problems in their sense of smell or taste. One location was on the Cornell University campus while the other was at a local shopping mall. Each consumer was presented with four samples of cottage cheese (about 30g per sample) that contained either 0.0, 0.25, 1.0, 4.0 ppm of

diacetyl in random order. Consumers rated five different attributes (appearance, odor, flavor, texture, and overall opinion) on 9-pt hedonic scales that were anchored with 'dislike very much' at one end and 'like very much' at the other. Space for written comments was provided. Consumers were also segmented by consumer group, based on frequency of consumption. "Frequent" consumers of cottage cheese were those who consumed cottage cheese at least two to three times per month (112 total) while "infrequent consumers" were those who consumed cottage cheese less than once per month (73 total).

Preparation of cottage cheese

Cottage cheese was made in the Cornell University pilot plant the week before each experiment began. A 5% inoculum of Miles Marschall Starter Culture OS (Marschall Products, Rhône-Poulenc, Inc., Madison, WI) was added to about 320 kg of pasteurized skim milk (71.7°C for 15 sec) in a Kusel L/I lab size model 8TH-2500 cottage cheese vat (Kusel, Inc., Watertown, WI) that was tempered to 32.2°C. When the pH reached 4.65 (about 4–5 hr after inoculation) the curd was cut with 1.27 cm knives, and cooked from 32.2°C to 54.4°C at a rate of about 5°C every 15 min. The curd was cooked an additional 18 min at 54.4°C to ensure destruction of all starter culture organisms. The whey was drained from the vat following cooking and the curd was triple washed with pasteurized reverse osmosis water. After the final wash the curd was allowed to drain for 30 min. Cream dressing (11.5% fat, 2.2% NaCl, and 2.4% NFDM) was then mixed with the curd in a ratio of 1.5 parts curd to 1.0 part dressing in the vat. The curd was then stored at 4.0°C overnight before being packaged into sealed ≈5L plastic containers. Fresh batches of cottage cheese were prepared for each experiment and were no more than 10 days old at completion of the experiment.

Diacetyl purification and addition

Diacetyl was purified as described by Meilgaard et al. (1982), with several modifications. Food grade diacetyl (Aldrich Chemical Co., Milwaukee, WI), 100 mL was distilled through a 100 cm long Vigreux column and collected in 10 mL fractions. Fractions 4–6 were transferred in 5 mL aliquots to an adsorption mixture that consisted of 15 mL glass centrifuge tubes that contained 2.76 g aluminum oxide (E. Merck Labs Inc. grade 90, neutral, activity grade I, EM Science, Inc., Gibbstown, NJ) and 0.65g activated carbon (Darco grade S-51 FF, American Norit Co., Inc., Jacksonville, FL) and was pre-dried at 105°C for 4 h. The slurry was vortexed and agitated in a 50.0°C water bath for 1 h and centrifuged at 6,000 × g for 30 min. The aqueous phase was removed and filtered through a pre-dried column (4 hr at 105°C) that consisted of a wide-mouthed Pasteur pipette with a glass wool plug that contained 0.55 g aluminum oxide (E. Merck Labs Inc. grade 90, neutral, activity grade I). The filtrate was collected in 1.5 mL micro-centrifuged tubes and micro-centrifuged for 10 min.

Stock solutions of 10,000 ppm and 3200 ppm diacetyl in spring water (Chemung Spring Water Co., Inc., Chemung, NY) were prepared and stored in amber bottles at 4°C. Calculated amounts of diacetyl were prepared each day from these stock solutions and were added quantitatively to a weighed amount of cottage cheese to give the desired concentration. Each dilution was prepared so that all samples received the same amount of liquid (1.0 g/100 g cheese). Those samples with no added diacetyl had added spring water. Samples were prepared 24 hr prior to their use and stored at 4°C to allow the diacetyl and the cottage cheese to reach equilibrium (Siebert, 1992). New stock solutions were prepared for each experiment.

Chemical and microbiological analyses

The level of diacetyl in each sample was determined before the beginning of each day's testing. Diacetyl was separated from cottage cheese using the steam distillation procedure described by Antinone (1990) with the exception that 50 g of cottage cheese was used instead of 20 g and the addition of lactic acid prior to distillation was omitted. The distillate was then injected into a Varian 6000 gas chromatograph equipped with an FID detector and a 2 m glass column containing 100/120 mesh Poropak® QS. Temperatures were: injector, 205, oven 180 and detector 250°C. The flow rates were nitrogen 25, hydrogen 25 and air 275 mL/min. Sample areas were compared to a standard curve prepared by distilling and chromatographing known concentrations of diacetyl in distilled water. Each sample was run in duplicate and the average area of three injections/sample was calculated. The standard curve was run in quadruplicate. The lower limit of detection for this method was 0.1 ppm.

Fat analyses were performed on each batch of cheese using the Mojonnier method and moisture using the atmospheric oven method (Richardson, 1985). These tests were performed the week after each experiment was concluded to ensure that the cheeses were at least 4% milk fat and 75 to 80% moisture.

Microbiological analyses were also performed on each day of testing. Cottage cheeses (11.0 g) were weighed into a stomacher bag to which 99.0 mL of diluent were added. The mixture was then placed in a Stomacher (UAC House, London, England) for 2 min and the samples were plated. Total aerobic, psychrotrophic, yeast and mold, and coliform counts were determined using the aerobic plate count method (Speck, 1984), psychrotrophic plate count method (Speck, 1984), the antibiotic yeast and mold count (Speck, 1984), and the Violet Red Bile Agar (VRBA) coliform count (Richardson, 1985). Lactic streptococci were also enumerated on the first day of each experiment using M-17 agar (Speck, 1984). This was done to ensure that no active residual starter culture was present that could potentially produce diacetyl during storage. All dilutions and diluents were prepared by standard methods (Speck, 1984).

Statistical analyses

Diacetyl detection. One estimate of diacetyl detection levels was the concentration at which correct responses were 50% above chance for each of the three methods (Brown et al., 1978; ASTM 1993a). A detection threshold is historically defined as the physical intensity level of a stimulus that is detected half the time. In order to correct for chance performance, Abbot's formula (Finney, 1971) has been widely used in psychology for decades, as follows: adjusted percent correct = (percent correct minus chance) / (1 minus chance). The threshold then becomes the adjusted 50% correct or 50% "above chance" (ASTM, 1993b). Adjusted percents correct were also converted to probits and a probit function fitted using the maximum likelihood estimation routine from the SYSTAT 5.2 NONLIN module, an example of which is given by Wilkinson et al. (1992, p. 437). Values of threshold were then interpolated from the fitted probit functions (Finney, 1971).

A second estimate of threshold was based on the Thurstonian/signal detection parameter, d', with threshold defined as $d' = 1.0$. Percents correct were converted to d' (Ennis, in press) to facilitate comparisons of the three methods and adjust for the inherent difficulty of the triangle procedure.

Consumer acceptance test. Repeated measures analysis of variance (Wilkinson et al., 1992) using SYSTAT 5.2 MGLH module was performed on each of the cottage cheese sample scores for each acceptance scale to determine if there was a difference ($P < 0.05$) in consumer acceptance. Analysis of variance was also used to examine the effects of location and consumption patterns of consumers. Comparisons of means were made using Tukey's HSD method for multiple comparisons (Ott, 1988). Data from 18 individuals (four from the campus location and 14 from the shopping mall location) were not used because of incomplete ballots.

RESULTS & DISCUSSION

Diacetyl detection

Using the criterion of 50% above chance performance (ASTM 1993b; Brown et al., 1978), the threshold for diacetyl was ≈ 0.2 ppm in the 2AFC procedure. Levels for diacetyl discrimination in cottage cheese were higher than those reported in other media and dependent on the testing method employed and the level of training given. Meilgaard et al. (1982) reported the diacetyl threshold in beer to be 0.07 to 0.15 μg/mL while Bennett et al. (1965) found the threshold for diacetyl in milk and cream to be 0.01 ppm. These thresholds may have been lower than those found in cottage cheese because of the nature of the product or the testing methods employed. Mean percents correct from the triangle, 3-AFC, and 2-AFC tests were compared (Fig. 1). The fitted probit function also predicted the 2-AFC threshold at 0.2 ppm, while the interpolated thresholds were 10 ppm for the 3-AFC data and above the experimental concentration range for the triangle procedure. Panelists were not able to discriminate well between samples that contained

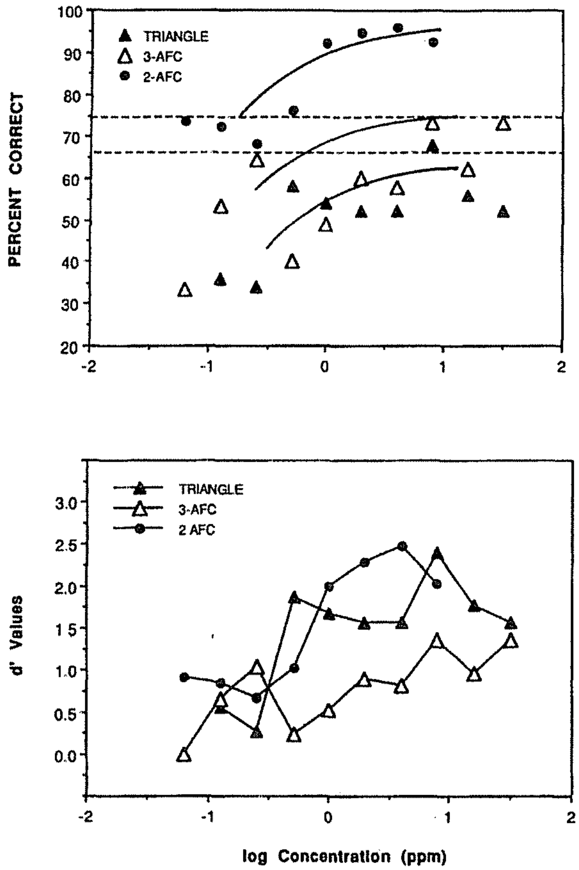

Fig. 1—Percentage of correct responses and d' values for diacetyl determination using three different discrimination tests methods. The upper and lower dotted lines represent 50% above chance performance for the two-product and three-product discrimination tests, respectively.

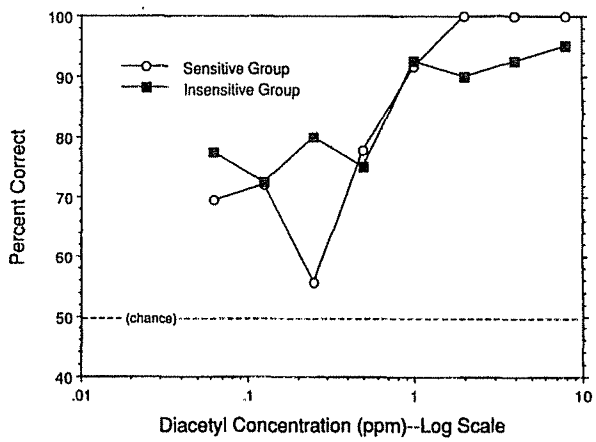

Fig. 2—Comparison of diacetyl sensitive and insensitive individuals performance using the 2-AFC test for diacetyl determination. Sensitive and insensitive individuals were selected based on their performance on the triangle and 3-AFC tests.

no added diacetyl and those with diacetyl using the triangle method and the 3 AFC procedure.

The criterion of 50%-above-chance is a useful benchmark, but not the only possible measure of discrimination. A Thurstonian/signal detection measure, d' (Frijters et al., 1980) or δ (Ennis, 1990; Ennis in press) can be derived which approximates an interval scale of perceptual distance for partially discriminable stimuli. Formally, d' estimates the separation of two distributions, one from the momentary sensations from the baseline sample and the other from the sensations from the augmented sample. The measure estimates this separation as distance between means in standard deviation units (Ennis, 1990). Figure 1 also shows the data replotted as d' values.

Choosing a d' value of 1.0 as an arbitrary threshold shows that the 2-AFC and triangle procedures emerge beyond this value in the range of about 0.2–0.5 ppm, while the 3-AFC data approach this level at much higher concentrations. Average d' values across all concentrations gives values of 1.53, 0.77, and 1.40 for the triangle, 3-AFC, and 2-AFC tests, respectively. From this perspective, the triangle and 2-AFC results agree more than they appear to using raw percent correct. This is because the d'-values take into account the difficulty inherent in triangle test comparisons. Correct triangle test performance entails estimation of subjective intensity *differences* among all three possible pairs in order to pick the odd sample, while AFC tests entail comparisons of perceived intensities (Frijters, 1979). Other factors beyond the simple univariate signal detection model should be considered; however, since the 3 AFC data indicate much less discriminability than the other methods and furthermore, assyptotic performance for the three methods differed. For example, average performance with the 4 and 8 ppm levels gives d' values of 1.06 for the 3-AFC, 1.98 for the triangle test, and 2.66 for the 2 AFC.

In the triangle procedure, performance improved only marginally at concentrations well above the aqueous threshold for diacetyl (Lawless et al., 1993; Bennett et al., 1965; Tuorila et al., 1981) and well above levels found in fermented dairy products (Veringa et al., 1984). Performance was lower when two samples in a set had added diacetyl than when only one was augmented. In the concentration range of 4.0 to 32.0 ppm 64.3% of the panelists identified the odd sample correctly when only one sample had added diacetyl, while only 48.6% identified the odd sample correctly when two samples were augmented. Similar results were described by O'Mahony and Goldstein (1986) in sweetened beverages who reported that correct performance was higher when the odd sample was sweetened than when the odd sample was unsweetened.

Ennis (1990) predicted that the 3-AFC procedure should be more powerful than the triangle test method. Frijters et al. (1982) suggested that individuals may convert the triangle test into a 3-AFC test in order to optimize the number of correct responses they give. However, foods are multidimensional. Some panelists may select an attribute that is not related to the attribute of interest, but which varies in a product such as cottage cheese (e.g. texture or appearance). Among panelists who gave incorrect answers for the cottage cheese triangle test 15% stated that texture and/or appearance was their main criterion for sample selection. This is not unreasonable since cottage cheese may have small discrepancies in appearance and texture among samples from the same lot. Such discrepancies may have distracted some panelists from differences in diacetyl flavor. Tourila et al. (1981) reported no differences between a triangle and a 2-AFC test for diacetyl odor threshold determination. However, since they performed threshold tests using diacetyl in water, it is unlikely that panelists used criteria other than aroma.

The 3-AFC method was employed to see if better results could be obtained when panelists were presented with only one augmented sample in each triad and were instructed to focus on differences in flavor. While there was improvement in the group threshold (approximately 3.0 ppm), there was only moderate improvement in discrimination at the higher diacetyl concentrations where a higher percent correct was expected. Several possible explanations exist for the lack of improvement for the 3-AFC method. The presence of a diacetyl insensitive segment

NYC-LAWLESS-000083

Table 1—Mean hedonic values for six different attributes of 4% milk fat cottage cheese

		Diacetyl Concentration							
		0.0 ppm		0.25 ppm		1.0 ppm		4.0 ppm	
Location	Attribute	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Locations Combined (N = 185)	Appearance	5.91^{a}	0.15	5.51^{a}	0.15	5.53^{a}	0.14	5.52^{a}	0.14
	Aroma	4.85^{a}	0.15	5.25^{ab}	0.15	5.58^{bc}	0.14	5.88^{c}	0.15
	Flavor	5.44^{a}	0.15	5.66^{ab}	0.16	6.14^{b}	0.15	5.99^{b}	0.15
	Texture	5.64^{a}	0.16	5.63^{a}	0.16	5.77^{a}	0.16	5.89^{a}	0.16
	Overall	5.64^{a}	0.14	5.65^{a}	0.14	5.98^{a}	0.14	5.95^{a}	0.14
Cornell Campus (N = 96)	Appearance	6.25^{a}	0.18	5.60^{b}	0.19	5.59^{b}	0.20	5.64^{b}	0.20
	Aroma	5.01^{a}	0.19	5.20^{ab}	0.19	5.71^{bc}	0.19	6.02^{c}	0.19
	Flavor	5.36^{a}	0.20	5.69^{a}	0.21	6.38^{b}	0.17	5.90^{ab}	0.20
	Texture	5.97^{a}	0.21	5.77^{a}	0.21	5.88^{a}	0.20	5.97^{a}	0.21
	Overall	5.43^{a}	0.18	5.50^{ab}	0.19	6.03^{b}	0.16	5.88^{ab}	0.19
Shopping Mall (N = 86)	Appearance	5.55^{a}	0.23	5.42^{a}	0.23	5.46^{a}	0.20	5.40^{a}	0.22
	Aroma	4.67^{a}	0.24	5.30^{ab}	0.23	5.44^{b}	0.21	5.73^{b}	0.24
	Flavor	5.52^{a}	0.23	5.64^{a}	0.24	5.88^{a}	0.25	6.09^{a}	0.24
	Texture	5.28^{a}	0.22	5.47^{a}	0.23	5.65^{a}	0.25	5.80^{a}	0.24
	Overall	5.86^{a}	0.21	5.81^{a}	0.22	5.93^{a}	0.24	6.02^{a}	0.21

^{a,b,c,d} Different superscripts indicate means within the rows that are significantly different ($P < 0.05$) using Tukey's HSD method for comparison of means.

(Lawless et al., in press; Meilgaard et al., 1982) could explain why performance only approached 75% correct even at 32.0 ppm. Another explanation may lie in the instructions given to panelists. Several panelists reported the presence of a sour flavor in the samples that did not contain diacetyl. Of the panelists who identified the incorrect sample in the 3-AFC test, an average of 5% reported that they based their selection on a higher sour note. Other panelists may have chosen samples based on higher sour notes as well. In preliminary experiments, samples of cottage cheese with moderate to high levels of diacetyl were rated less sour than those with lower levels of diacetyl (Antinone et al., 1990). Diacetyl may mask or counteract sour flavor for some individuals similar to the effect of odor counteraction (Cain and Drexler, 1974). Hence, panelists may have chosen an incorrect sample because they were instructed to choose a sample with the strongest flavor, not one lacking in flavor. Fatigue and/or adaption (Byer and Abrams, 1953; Meiselman, 1972) may have also contributed to the poor performance in the 8 to 32 ppm concentration range, since all 3 of the highest concentrations were evaluated in the same session.

The 2-AFC test was performed to attempt to maximize discrimination and it included a training session. Aqueous samples of diacetyl (0.5 and 4.0 ppm) were presented to panelists to familiarize them with diacetyl aroma. Panelists were asked to smell and taste the samples and then discuss the different flavor characteristics associated with the compound. Panelists were then give three samples of cottage cheese that contained either 0.0, 0.5, or 4.0 ppm diacetyl. Since panelists were told the diacetyl concentration in each sample, they were able to define their own criteria for what diacetyl "tastes" like. Hence, whether an individual detected the buttery note of diacetyl or noticed a decrease in sour flavor, they would be able to correctly identify the sample that contained the added diacetyl. The 2-AFC procedure was also chosen to minimize fatigue and adaptation by presenting fewer samples (Byer and Abrams, 1953). The improvement with the 2 AFC procedure may have been due to training the panelists, less fatigue and adaptation due to fewer samples; the nature of the discrimination process (Frijters, 1979; Frijters et al., 1982) or a combination of these factors.

The difficulties encountered in interpreting results of the three types of difference tests suggest that researchers studying flavor thresholds should carefully consider the testing method. In addition, the use of the 50%-above chance criterion for threshold may fail to identify differences among samples, as in the case of our triangle data. The criterion of 50% above chance in the triangle procedure requires a d' value of 2.3, which is a very high level of discriminability from the perspective of signal detection analysis.

Since previous work showed individual differences in diacetyl sensitivity (Lawless et al., in press), the performance of sensitive and insensitive groups in diacetyl detection were compared in the 2 AFC phase (Fig. 2). While the sensitive segment performed better at higher concentrations, very little overall difference was observed between the two groups. Thus, either diacetyl insensitivity was not a factor in the poor performance for the triangle and 3-AFC methods, or diacetyl sensitivity may be a transient phenomenon in which individuals vary over time (Stevens et al., 1988).

Consumer acceptance test

The data for the consumer acceptance tests of diacetyl in full fat cottage cheese were compared (Table 1). Mean aroma ratings increased with increasing diacetyl concentration. The 4.0 ppm sample was different from the 0.0 and 0.25 ppm samples and the 1.0 ppm sample was different from the 0.0 ppm sample. Mean flavor ratings showed a similar increase with concentration, but peaked at 1.0 ppm then declined at 4.0 ppm. Both the 1.0 and 4.0 ppm samples were different from the 0.0 ppm sample. Mean overall opinion ratings showed a similar pattern to those for flavor, but no statistical differences. Consumers may have weighted attributes other than diacetyl flavor in judging overall appeal. For example, 6% of the panelists commented that samples were runny or watery. Sample dryness and non-uniform curd size were also mentioned. Appearance and texture showed relative uniformity amongst samples with the exception of a high appearance rating for the 0.0 ppm sample in the campus group.

No effect of location or interaction of location with diacetyl levels was observed for any rated attributes. Both groups showed an increase in mean hedonic ratings for aroma, flavor, and overall opinion with increasing diacetyl concentration with the exception of flavor ratings for the campus group which peaked at 1.0 ppm. However, the mall group was less discriminating (Table 1). Standard deviations of the mall group were higher by an average of 0.273 scale units (Mann-Whitney U-Test, $P < 0.05$). Environmental effects may have been responsible for the decrease in discrimination for the mall group. The test area in the shopping mall was noisy, crowded, and not odor free. The campus testing was done in a well lit, odor free, and quiet environment that was designed for sensory testing. Claassen (1991) found little difference between campus groups and random groups from the surrounding community in perception of flavor defects in milk. Both of those groups, however, performed the tests in a controlled sensory testing facility.

Effects of consumption patterns on acceptance of diacetyl was examined. The infrequent consumers rated the 0.0 ppm sample higher than the other three on the appearance attribute

($P < 0.05$). Frequent consumers also rated the 0.0 ppm sample higher than the other three, but to a lesser extent. P-values for main effects of consumption frequency were: flavor acceptance 0.07, aroma acceptance 0.07, and overall opinion 0.06, consistent with slightly higher acceptance among frequent consumers. The infrequent consumers rated all samples, except the 0.0 ppm sample, lower for texture than the frequent consumer group. No interaction of consumption group with diacetyl level was observed.

CONCLUSIONS

THE THRESHOLD for diacetyl, 0.2 ppm, was higher than has been reported in other media but still low enough for addition of diacetyl to cottage cheese to be practical. Cottage cheese with 1.0 ppm of diacetyl was maximally accepted for combined flavor and aroma appeal by consumers. These data were contradictory to those previously reported using dairy judges, but were in agreement with the recent results for consumers. For the detection of diacetyl in cottage cheese, the two-alternative forced choice method coupled with a brief familiarization with diacetyl was more sensitive than the three alternative forced choice procedure without familiarization. Giving specific instructions about the nature of the product differences may help panelists discriminate.

REFERENCES

Antinone, M.J. 1990. Diacetyl in cottage cheese: its presence, determination, and microbial reduction by several spoilage bacteria. M.S. thesis, Cornell University, Ithaca, New York.

Antinone, M.J., Lawless, H.T., Ledford, R.A., Glatter, S. 1990. Unpublished data. Dept. of Food Science, Cornell University, Ithaca, NY.

ASTM 1993a. Standard practice for defining and calculating individual and group thresholds from forced-choice data sets in intermediate size (E-1432). *Annual Book of Standards* 15.07: 72–79.

ASTM 1993b. Standard terminology relating to sensory evaluation of materials and products (E-253). *Annual Book of Standards* 15.07: 1–4.

Bennett, G., Liska, B.J., and Hempenius, W.L. 1965. Effect of other flavor components on the perception of diacetyl in fermented dairy products. J. Food Sci. 30: 35–37.

Brown, D.G.W., Clapperton, J.F., Meilgaard, M.C., and Moll, M. 1978. Flavor thresholds of added substances. Am. Soc. Brew. Chem. J. 36: 73–80.

Byer, A.J., and Abrams, D. 1953. A comparison of the triangular and two-sample taste-test methods. Food Technol. 7: 185–187.

Cain, W.S., and Drexler, M. 1974. Scope and evaluation of odor counteraction and masking. Ann. New York Acad. Sci. 237: 427–439.

Claassen, M.R. 1991. A comparison of descriptive terminology systems for the sensory analysis of flavor defects in milk. M.S. thesis, Cornell Univ., Ithaca, New York.

Collins, E.B. 1972. Biosynthesis of flavor compounds by microorganisms. J. Dairy Sci. 55: 1022–1028.

Ennis, D.M. In press. The power of sensory discrimination methods. J. Sensory Studies.

Ennis, D.M. 1990. Relative power of difference testing methods in sensory evaluation. Food Technol. 44(4): 114–117.

Ernstrom, C.A. 1965. Continuous process developed for making cottage cheese curd without culture. Can. Dairy Ice Cream J. 44: 32–36.

Finney, D.J. 1971. *Probit Analysis*, 3rd ed. Cambridge Univ. Press, London.

Frijters, J.E.R. 1979. The paradox of discriminatory nondiscriminators resolved. Chem Senses Flav. 4: 355–358.

Frijters, J.E.R., Kooistra, A., and Vereijken, P.F.G. 1980. Tables for d' for the triangular method and the 3-AFC signal detection procedure. Percep. Psychophys. 27: 176–178.

Frijters, J.E.R., Blauw, Y.H., and Vermaat, S.H. 1982. Incidental training in the triangular method. Chem. Senses 7: 63–69.

Geilman, W. 1992. Consumer preference for cottage cheese with starter distillate.*California Dairy Foods Research Center 1991/92 Annual Report*, p. 23.

Lawless, H.T., Antinone, M.J., Ledford, R.A., and Johnston, M. In press. Olfactory responsiveness to diacetyl. J. Sensory Studies.

McBride, R.L. and Hall, C. 1979. Cheese grading versus consumer acceptability: an inevitable discrepancy. Aust. J. Dairy Technol. 34: 66–68.

Meilgaard, M.C., Reid, D.S., and Wyborski, K.A. 1982. Reference standards for beer flavor terminology system. Am. Soc. Brew. Chem. J. 40(4): 119–128.

Meiselman, H.L. 1972. Human taste perception. CRC Critical Reviews in Food Tech. 3: 89–119.

O'Mahony, M. and Goldstein, L.R. 1986. Effectiveness of sensory difference tests: Sequential sensitivity analysis for liquid food stimuli. J. Food Sci. 51: 1550–1553.

Ott, L. 1988. *An Introduction to Statistical Methods and Data Analysis*. PNS-Kent Publishing Co., Boston, MA.

Parker, R.B., and Elliker, P.R. 1963. Effect of spoilage bacteria on biacetyl content and flavor of cottage cheese. J. Dairy Sci. 36: 843–849.

Richardson, G.H., ed. 1985. *Standard Methods for the Examination of Dairy Products*, 15th ed. American Public Health Assoc., Washington, DC.

Schroeder, C.L., Bodyfelt, F.W., Wyatt, C.J., and McDaniel, M.R. 1988. Reduction of sodium chloride in cheddar cheese: effect on sensory, microbiological, and chemical properties. J. Dairy Sci. 71: 2010–2020.

Sebek, O.K., and Randles, C.I. 1952. The oxidation of the steroisomeric 2,3-butanediols by *Pseudomonas*. Arch. Biochem. Biophys. 40: 373–380.

Siebert, K. 1992. Personal communication. Dept. of Food Science Geneva campus, Cornell University, Ithaca, NY.

Speck, M.L., ed. 1984. *Compendium of Methods for the Microbiological Examination of Foods*. American Public Health Assoc., Washington, DC.

Stevens, J.C., Cain, W.S., and Burke, R.J. 1988. Variability of olfactory thresholds. Chem. Senses 13: 643–653.

Tuorila, H., Kurkela, R., Suihko, M., and Suhonen, U. 1981. The influence of tests and panelists on odor detection thresholds. Lebensm. Wiss. Technol. 15: 97–101.

Veringa, H.A., Verburg, E.H., and Stadhouders, J. 1984. Determination of diacetyl in dairy products containing alpha-acetolactic acid. Neth. Milk Dairy J. 38: 251–263.

Wilkinson, L., Hill, M. and Vang, E. 1992. SYSTAT, *Statistics*. SYSTAT, Inc., Evanston, IL.

Ms received 2/25/93; revised 8/3/93; accepted 11/9/93.

We thank L. Mattick for help with GC analysis, R. Vaia, M. Cornwall-Brady, and M. Mattison for help with testing procedures, and C. Scinto, E. Halstead, and J. Davidson for help in cottage cheese manufacturing. We thank Dr. D. Ennis for providing tables for d'. This work was funded by a grant from the Northeast Dairy Foods Research Center.